Necessary optimality conditions for abnormal abstract problems with general constraints

Aram Arutyunov*, Fernando Lobo Pereira**

* Nonlinear Analysis and Optimization Dept.,
Peoples Friendship University of Russia
6, Mikluka-Maklai St., Moscow, 117198, Russia,
email: arutun@orc.ru

** Instituto de Sistemas e Robótica,
Faculdade de Engenharia da Universidade do Porto,
R. Dr. Roberto Frias, 4200-465 Porto, Portugal,
email: flp@fe.up.pt

Abstract

Here, we derive second-order necessary conditions of optimality for an abstract optimization problem with equality and inequality constraints and constraints in the form of an inclusion into a given closed set. An important feature is that our optimality conditions dispense with any a priori normality assumptions, such as Robinson's constraint qualification, and remain informative even for abnormal points.

Problem Statement

Given a vector space X, a set $C \subseteq X$, mappings $F_1 : X \to \mathbb{R}^{k_1}$, $F_2 : X \to \mathbb{R}^{k_2}$, and a function $f : X \to \mathbb{R}^1$, we consider the following optimization problem

$$(P) \qquad f(x) \to \min, \quad x \in C, \quad F_1(x) \leq 0, \quad F_2(x) = 0. \qquad (1)$$

Our main goal is to obtain first- and second-order necessary extremum conditions for this problem under some assumptions about smoothness and properties of the set C. Let us introduce the notation and assumptions needed in what follows.

Let $k = k_1 + k_2$ and $F = (F_1, F_2) : X \to Y = \mathbb{R}^k$. We consider the so-called finite topology in the vector space X. We denote by $\mathcal{M}$ the set of all finite-dimensional linear subspaces $M \subseteq X$. A set is open in the finite topology if and only if its intersection with any $M \in \mathcal{M}$ is open in the unique Hausdorff vector topology of M[1]. We denote the finite topology by τ. If a space X is infinite-dimensional, then, equipping it with the finite topology τ, does not, in general, turn it into a topological vector space because the addition is, as a rule, discontinuous. On the other hand, the finite topology is stronger than all topologies that transform X into a vector topological space and, therefore, a local minimum with respect to the finite topology is the weakest one amongst all types of minima considered.

Fix a point $x_0 \in X$. We assume that functions f and F are twice continuously differentiable in a neighborhood of x_0 with respect to the finite topology τ. This means that, for an arbitrary $M \in \mathcal{M}$ containing the point x_0, the restrictions of f and F to M are twice continuously differentiable in a certain (depending on M) neighborhood of x_0. Therefore, there exist a linear functional $a \in X^*$ (X^* is the space algebraically dual to X), a linear operator $A : X \to Y$, a bilinear form $q : X \times X \to \mathbb{R}^1$, a bilinear mapping $Q : X \times X \to Y$, and mappings $\alpha_0 : X \to \mathbb{R}^1$, and $\alpha : X \to Y$, such that, $\forall x \in X$

$$
\begin{aligned}
f(x) &= f(x_0) + \langle a, x - x_0 \rangle + \frac{1}{2} q[x - x_0]^2 + \alpha_0(x - x_0), \\
F(x) &= F(x_0) + A(x - x_0) + \frac{1}{2} Q[x - x_0]^2 + \alpha(x - x_0),
\end{aligned}
$$

and, for an arbitrary $M \in \mathcal{M}$, such that $x \in M$,

$$
\frac{\alpha_0(x - x_0)}{\|x - x_0\|_M^2} \to 0, \text{ and } \frac{|\alpha(x - x_0)|}{\|x - x_0\|_M^2} \to 0, \text{ as } x \to x_0,
$$

where $\|\cdot\|_M$ is a finite-dimensional norm in M.

Here and in what follows, we denote a bilinear mapping B by $B[x,x]$ or $B[x]^2$. The mappings a and q are denoted, respectively, by $\frac{\partial f}{\partial x}(x_0)$ and $\frac{\partial^2 f}{\partial x^2}(x_0)$. They are called the first- and second-order derivatives of f. A similar notation is used for the mapping F and other functions.

Regarding the set C, we assume it is closed in the finite topology.

The first-order necessary conditions for problem (P) used below are based on Mordukhovich's normal cone to the set C at $x_0 \in C$. Take any linear subspace $M \in \mathcal{M}$ such that $x_0 \in C \cap M$. Let us introduce Mordukhovich's normal cone [6,8] to the set $C \cap M$ at x_0 denoted by $N^M(x_0, C)$. For $x \in M$, we put $W(x, M \cap C) = \{w \in M \cap C : \|x - w\|_M = dist_M(x, C)\}$, where $dist_M(x, C) = \inf_{\xi \in C \cap M} \{\|\xi - x\|_M\}$, denotes the distance function.

Let cone$A$ and cl$A$ denote, respectively, the conic hull and the closure of the set A. Consider the upper semicontinuous hull of the set

$$
\text{cone}(x - W(x, M \cap C)) = \bigcup_{r > 0} \Big\{ r(x - W(x, M \cap C)) \Big\}
$$

[1] Note that any finite-dimensional vector space is equipped with a unique Hausdorff vector topology.

as $x \to x_0, x \in M$:

$$\begin{aligned} N^M(x_0,C) &= \bigcap_{\varepsilon>0} \mathrm{cl}\Big[\bigcup_{\substack{x\in M\\ \|x-x_0\|_M\le\varepsilon}} cone(x-W(x,M\cap C))\Big] \\ &:= \limsup_{\substack{x\in M\\ x\to x_0}} cone(x-W(x,M\cap C)). \end{aligned}$$

We put

$$N(x_0,C) = \bigcup_{M\in\mathcal{M}} N^M(x_0,C).$$

Mordukhovich's normal cone $N(x_0,C)$ is the smallest and the most natural one to derive necessary conditions of optimality, [8,9,11]. Indeed, $N^M(x,C)$ is upper semicontinuous in $C\cap M$ and, if x_0 is the solution to the problem of minimizing a smooth function f over a set C, then $\frac{\partial f}{\partial x}(x_0) \in -N(x_0,C)$ (see [8, 9]). Furthermore, $N(x_0,C)$ is closed but may be nonconvex. At the same time, the closure of its convex hull coincides with Clarke's normal cone (see [8, 9]).

We consider the second-order variations of the set C at a point $x \in C$ in a direction d (see [4, 10]). That is,

$$\begin{aligned} T^2_{C\cap M}(x,d) &= \Big\{w\in X : dist_M(x+\varepsilon d+\frac{1}{2}\varepsilon^2 w,C) = o(\varepsilon^2),\, \varepsilon\ge 0\Big\}, \\ O^2_{C\cap M}(x,d) &= \Big\{w\in X : \exists\{\varepsilon_n\}\downarrow 0 \text{ s.t. } dist_M(x+\varepsilon_n d+\frac{1}{2}\varepsilon_n^2 w,C) = o(\varepsilon_n^2)\Big\}, \end{aligned}$$

where $M\in\mathcal{M}$ is an arbitrary finite-dimensional linear subspace containing x and d.

Sets $T^2_{C\cap M}(x,d)$ and $O^2_{C\cap M}(x,d)$ are called inner and outer second order tangent sets, respectively (see [4, 10]). It was pointed out in this reference that both these sets are closed, that $T^2_{C\cap M}(x,d)\subseteq O^2_{C\cap M}(x,d)$, and that $O^2_{C\cap M}(x,d)\neq\emptyset$ only if $d\in T_{C\cap M}(x)$. Here, as usual,

$$T_{C\cap M}(x) = \{d\in M : \exists\varepsilon_n\downarrow 0,\, dist_M(x+\varepsilon_n d, C\cap M) = o(\varepsilon_n)\}$$

is the contingent (Bouligand) cone to the set $C\cap M$ at the point x. Also, if the set C is convex, then the inner tangent set $T^2_{C\cap M}(x,d)$ is convex, but the outer tangent set $O^2_{C\cap M}(x,d)$ may be nonconvex, [4, 10]. We put

$$T^2_C(x,d)=\bigcup_{M\in\mathcal{M}} T^2_{C\cap M}(x,d),\ O^2_C(x,d)=\bigcup_{M\in\mathcal{M}} O^2_{C\cap M}(x,d),\ T_C(x)=\bigcup_{M\in\mathcal{M}} T_{C\cap M}(x).$$

It is evident that if the set C is convex, then the inner tangent set $T^2_C(x,d)$ is also convex. The computation of second-order tangent sets for positive cones in some specific spaces may be found in [5].

A linear subspace $\mathcal{I}\subseteq X$ such that $x+\mathcal{I}\subseteq C$, $\forall x\in C$ is named invariant linear subspace (ILS) with respect to C.

For a given closed set C, an ILS, in general, is nonunique. Clearly, any linear subspace of an ILS also is an invariant linear subspace. For example, $\{0\}$ is ILS relatively to any closed set C.

For any $x \in C$, put $\mathcal{I}_C(x) = \bigcap\limits_{r \neq 0} r(C - x)$, where the intersection is taken over all reals $r \in \mathbb{R}^1$, $r \neq 0$.

Main result

Consider the cone $\mathcal{K} := \mathcal{K}(x_0)$ defined by

$$\mathcal{K}(x_0) \quad := \quad \Big\{ h \in T_C(x_0) : \, \langle \frac{\partial f}{\partial x}(x_0), h \rangle \leq 0, \; \frac{\partial F_2}{\partial x}(x_0) h = 0, \text{and} \\ \frac{\partial F_{1,j}}{\partial x}(x_0) h \leq 0 \; \forall j \text{ such that } F_{1,j}(x_0) = 0 \Big\}.$$

Here, $F_{1,j}$ are the coordinates of the vector function F_1. The cone $\mathcal{K}$ is, obviously nonempty (because it contains zero) and convex if the set C is convex. The cone $\mathcal{K}(x_0)$ is called the critical cone of the problem (P) at the point x_0.

Let $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ with $\lambda_0 \in \mathbb{R}^1$, $\lambda_1 \in \mathbb{R}^{k_1}$, and $\lambda_2 \in \mathbb{R}^{k_2}$, and consider the (generalized) Lagrangian

$$\mathcal{L}(x, \lambda) = \lambda_0 f(x) + \langle \lambda_1, F_1(x) \rangle + \langle \lambda_2, F_2(x) \rangle.$$

Let $\Lambda = \Lambda(x_0)$ denote the set of the generalized Lagrange multipliers λ that correspond to the point x_0 according to the Lagrange multiplier rule (also known as Fritz John optimality conditions)(see [4, 6, 8]), i.e., such that

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial x}(x_0, \lambda) \in -N(x_0, C) \\ \lambda_0 \geq 0, \lambda_1 \geq 0, \, \langle \lambda_1, F_1(x_0) \rangle = 0, \, |\lambda| = 1. \end{cases} \tag{2}$$

In what follows, we assume, for convenience, that $F_1(x_0) = 0$. This can always be achieved by omitting the nonactive components of the inequality constraints.

Take any linear subspace $M \subseteq X$ and consider the set of all Lagrange multipliers $\lambda \in \Lambda$ for which there exists a linear subspace $\Pi \subseteq M$ (depending on λ) such that

$$\operatorname{codim}_M \Pi \leq k, \quad \Pi \subseteq \operatorname{Ker} \frac{\partial F}{\partial x}(x_0), \quad \frac{\partial^2 \mathcal{L}}{\partial x^2}(x_0, \lambda)[x, x] \geq 0, \quad \forall x \in \Pi. \tag{3}$$

Here and in what follows, codim_M denotes the codimension relative to the subspace M. We denote the set of such Lagrange multipliers by $\Lambda(x_0, M)$. Each set $\Lambda(x_0, M)$ is obviously compact (but it may be empty).

For a given set $T \subseteq X$, we denote by $\sigma(\cdot, T)$ its support function, i.e., $\sigma(x^*, T) = \sup\limits_{x \in T} \langle x^*, x \rangle$, $x^* \in X^*$. (Extremum principle). Let x_0 be a point of local minimum with respect to the finite topology τ of the problem under consideration.

Then, for each ILS $\mathcal{I}$ with respect to C, the set $\Lambda(x_0,\mathcal{I})$ is nonempty, and, moreover, for each $h \in \mathcal{K}(x_0)$ and any convex set $\mathcal{T}(h) \subseteq O^2_C(x_0,h)$,

$$\max_{\lambda\in\Lambda_a}\left(\frac{\partial^2\mathcal{L}}{\partial x^2}(x_0,\lambda)[h,h]-\sigma\left(-\frac{\partial\mathcal{L}}{\partial x}(x_0,\lambda),\mathcal{T}(h)\right)\right)\geq 0. \tag{4}$$

Here, $\Lambda_a = \mathrm{conv}\Lambda(x_0,\mathcal{I})$ and conv denotes the convex hull of a set.

For the case $C = X$, theorem was obtained in [1]. If C is a convex cone[2], then, for $h \in C + span\{x_0\}$, theorem was obtained in [2, 3].

Proof of the Extremum Principle.

The proof of theorem is based on the removing the constraints $F_1(x) \leq 0$ and $F_2(x) = 0$ by using the penalty method. A central role in the proof is played by the following assertion that allows us to pass to the limit.

Let X be a Banach space and $\{\Pi_n\}$ be a sequence of closed linear subspaces of X such that $\mathrm{codim}\Pi_n \leq k$, $\forall n$.

Then, there exists a closed linear subspace $\Pi \subseteq X$ such that

$$\mathrm{codim}\Pi \leq k, \quad \Pi \subseteq \mathrm{Ls}\{\Pi_n\}.$$

Here and in what follows, Ls is the upper topological limit of a sequence of sets.[3]

Theorem is proved in §1.7 of [1]. We need its following modification which is, in fact, equivalent to the theorem.

Let $A_n : X \to \mathbb{R}^k$ be a sequence of continuous linear operators which converges by the norm to a linear operator A.

Then, there exists a linear closed subspace $\Pi \subseteq X$ such that

$$\mathrm{codim}\Pi \leq k, \quad \Pi \subseteq \mathrm{Ls}\{\mathrm{Ker}A_n\}, \quad \Pi \subseteq \mathrm{Ker}A.$$

Theorem is proved by setting $\Pi_n = \mathrm{Ker}A_n$ in theorem .

Proof. We divide the proof of theorem into four stages. First, in stage I, we prove that the set Λ_a is nonempty for the case in which the space X is assumed to be finite-dimensional.

Then, we show, by using the results obtained above, that, in the case where X is finite-dimensional, for each $h \in \mathcal{K}(x_0)$, and $w \in O^2_C(x_0,h)$, $\exists\, \lambda \in \Lambda_a$ such that

$$\frac{\partial^2\mathcal{L}}{\partial x^2}(x_0,\lambda)[h,h]+\langle\frac{\partial\mathcal{L}}{\partial x}(x_0,\lambda),w\rangle\geq 0. \tag{5}$$

In stage III, we prove (5) in its full generality, i.e., we omit the assumption $\dim X < \infty$. Finally, in stage IV, we prove (4).

[2] Note that if C is a convex cone, then the tangent cone $T_C(x)$ coincides with closure of the set $C + span\{x\}$ for each $x \in C$.

[3] Note that the upper topological limit of a sequence of sets $\{\Pi_n\}$ is the set of all limit points of the sequences $\{x_n\}$ such that $x_n \in \Pi_n$, $\forall n$.

STAGE I. We assume that X is finite-dimensional. By using this assumption, we define the inner product in X and transform it into an Euclidean space. Let us remove the constraints $F_1 \le 0$, $F_2 = 0$ in the problem under consideration by the penalty method. For each positive integer i, we set

$$f_i(x) = f(x) + i\Big(\sum_{j=1}^{k_1} (F_{1,j}^{+}(x))^4 + |F_2(x)|^4\Big) + |x - x_0|^4,$$

where $a^+ = \max\{a, 0\}$, and consider the following family of minimization problems, called i-problems,

$$f_i(x) \to \min, \quad x \in C, \quad |x - x_0| \le \delta.$$

Here, $\delta > 0$ is chosen in such a way that x_0 is a minimum for the initial problem in a δ-neighborhood of the point x_0 (recall that the space X is assumed to be finite-dimensional at this stage, and, therefore, this δ does exist). The solution to the i-problem is denoted by x_i.

We prove that $x_i \to x_0$, as $i \to \infty$. Indeed, by taking into account that X is finite-dimensional and, by extracting a subsequence (if necessary), we obtain the convergence of the sequence $\{x_i\}$ to a certain $\bar{x}$.

Now, we show that $\bar{x} = x_0$. Indeed, the fact $f_i(x_i) \le f_i(x_0) = f(x_0)$, $\forall i$, implies that

$$\overline{\lim_{i\to\infty}} F_{1,j}(x_i) \le 0, \ \forall j, \text{ and } F_2(x_i) \to 0, \text{ as } i \to \infty. \tag{6}$$

This implies that $F_1(\bar{x}) \le 0$ and $F_2(\bar{x}) = 0$. In addition, the first inequality in (6) implies that $f(x_i) + |x_i - x_0|^4 \le f(x_0)$, $\forall i$, and, therefore, that $f(\bar{x}) + |\bar{x} - x_0|^4 \le f(x_0) \le f(\bar{x})$. Thus, $\bar{x} = x_0$.

For large i's, which is our sole concern, we have $|x_i - x_0| < \delta$ and, consequently, the constraint $|x - x_0| \le \delta$ is not active. In another words, the i-problem is locally a finite-dimensional problem with the constraint $x \in C$. The first-order necessary condition (see [7]) for this problem gives

$$\frac{\partial f_i}{\partial x}(x_i) \in -N(x_i, C). \tag{7}$$

Let us prove the second-order necessary condition for this problem

$$\frac{\partial^2 f_i}{\partial x^2}(x_i)[x,x] \ge 0, \quad \forall x \in \mathfrak{I}. \tag{8}$$

Indeed, by virtue of the definition of ILS, we have

$$x_i + \varepsilon x \in C, \quad \forall x \in \mathfrak{I}, \ \forall \varepsilon. \tag{9}$$

Define the scalar function ϕ as follows $\phi(\varepsilon) = f_i(x_i + \varepsilon x)$. This function ϕ is smooth and in virtue of (9) it attains the local minimum at $\varepsilon = 0$. Consequently, $\phi''(0) \ge 0$. Calculating the second derivative of ϕ we obtain $\phi''(0) = \frac{\partial^2 f_i}{\partial x^2}(x_i)[x,x] \ge 0$, which proves (8).

Now, we decode conditions (7), (8) in terms of the data of the problem. Let $\bar{\lambda}_{1,i}^{j} = 4iF_{1,j}^{+}(x_i)|F_{1,j}^{+}(x_i)|^3$, $\bar{\lambda}_{1,i} = (\bar{\lambda}_{1,i}^{1},\ldots,\bar{\lambda}_{1,i}^{k_1})$, $\bar{\lambda}_{2,i} = 4iF_2(x_i)|F_2(x_i)|^3$, $\lambda_{0,i} = (1+|\bar{\lambda}_{1,i}|^2 + |\bar{\lambda}_{2,i}|^2)^{-\frac{1}{2}}$, $\lambda_{1,i} = \lambda_{0,i}\bar{\lambda}_{1,i}$, $\lambda_{2,i} = \lambda_{0,i}\bar{\lambda}_{2,i}$, and $\lambda_i = (\lambda_{0,i},\lambda_{1,i},\lambda_{2,i})$.

Then, conditions (7) and (8) take the form, respectively,

$$\frac{\partial \mathcal{L}}{\partial x}(x_i,\lambda_i) + 1(i) \in -N(x_i,C), \tag{10}$$

$$\frac{\partial^2 \mathcal{L}}{\partial x^2}(x_i,\lambda_i)[x,x] + 12i\lambda_{0,i}\left(|F_1^{+}(x_i)|^2\Big|\frac{\partial F_1}{\partial x}(x_i)x\Big|^2 + |F_2(x_i)|^2\Big|\frac{\partial F_2}{\partial x}(x_i)x\Big|^2 + 1(i)\right) \geq 0, \quad \forall x \in \mathfrak{I}, \text{ s.t. } |x| \leq 1. \tag{11}$$

Here, $1(i)$ is a sequence converging to zero. By construction,

$$|\lambda_i| = 1,\ \lambda_{0,i} \geq 0,\ \lambda_{1,i} \geq 0,\ \forall i. \tag{12}$$

By extracting a subsequence, we obtain a $\lambda = (\lambda_0,\lambda_1,\lambda_2)$ such that $\lambda_i \to \lambda$, $i \to \infty$. By passing onto the limit in (10) and (12), and by using the upper semicontinuity over the set C of the set-valued mapping $N(\cdot,C)$, we conclude that $\lambda \in \Lambda$.

Let us prove the existence of a linear subspace Π, satisfying (3). For this purpose, we define linear operators $A_i : \mathfrak{I} \to \mathbb{R}^k$ by the formula $A_i x = F'(x_i)x$. By theorem , there exists a linear subspace $\Pi \subseteq \mathfrak{I}$ such that

$$\Pi \subseteq \operatorname{Ker} F'(x_0), \quad \Pi \subseteq \operatorname{Ls}\{\operatorname{Ker} A_i\}, \quad \operatorname{codim}_{\mathfrak{I}} \Pi \leq k.$$

Take an arbitrary vector $h \in \Pi$. By the definition of the upper topological limit, there exists a sequence $\{h_i\}$, with $h_i \in \mathfrak{I} \bigcap \operatorname{Ker} F'(x_i)$, such that, by extracting a subsequence, we obtain $h_i \to h$ as $i \to \infty$. By substituting $x = h_i$ into (11) and by passing onto the limit we obtain (3). Therefore, $\lambda \in \Lambda_a$ and, thus, $\Lambda_a \neq \emptyset$.

STAGE II. As in the previous stage, we consider X to be finite-dimensional. For convenience, we assume that $f(x_0) = 0$. Fix an arbitrary vector $h \in \mathcal{K}(x_0)$ with $|h| = 1$ and take any $w \in O_C^2(x_0,h)$. Let us prove (5).

By the definition there exists a sequence $\{\varepsilon_n\} \downarrow 0$ such that

$$x_n = x_0 + \varepsilon_n h + \frac{1}{2}\varepsilon_n^2 w + o(\varepsilon_n^2) \in C, \quad \forall \varepsilon_n. \tag{13}$$

Consider the function defined by

$$\gamma(\chi) = \begin{cases} 0 & \forall \chi \leq 1 \\ (\chi-1)^4 & \forall \chi > 1. \end{cases}$$

For every positive integer n, consider the following minimization problem with the respect to the variables $(x,\chi) \in X \times \mathbb{R}^1$:

$$\begin{array}{ll} \text{Minimize} & f_n(x,\chi) \\ \text{subject to:} & (x,\chi) \in C \times \mathbb{R}^1, \\ & F_1(x) - \chi F_1(x_n) \leq 0, \\ & F_2(x) - \chi F_2(x_n) = 0, \\ & 0 \leq \chi, \text{ and } |x-x_0|^2 \leq \delta^2. \end{array} \tag{14}$$

This problem is called the $n-$problem. Here, $\delta > 0$ is defined as above and

$$f_n(x,\chi) = \tilde{f}(x) - \chi\tilde{f}(x_n) + \gamma(\chi) \text{ where } \tilde{f}(x) = f(x) + |x - x_0|^4.$$

For all n sufficiently large, we have $|\varepsilon_n h + \frac{\varepsilon_n^2}{2} w| < \delta$. Take any such n. There exists a solution to the n-problem, because the point (x,χ) defined by $x = x_n$ and $\chi = 1$ satisfies all the constraints of this problem, the ball $\{x : |x| \leq \delta\}$ is compact, and

$$\gamma(\chi)\chi^{-1} \to \infty \text{ as } \chi \to \infty. \tag{15}$$

We show that, among the solutions to the $n-$problem, there exists at least one $(\hat{x}_n, \hat{\chi}_n)$ such that $\hat{\chi}_n > 0$. Indeed, let $(\tilde{x}_n, \tilde{\chi}_n)$ be a solution to the $n-$problem. If $\tilde{\chi}_n > 0$ then we put $\hat{x}_n = \tilde{x}_n, \hat{\chi}_n = \tilde{\chi}_n$. Suppose that $\tilde{\chi}_n = 0$. Then $\tilde{x}_n$ is feasible in problem (P). If $\tilde{x}_n \neq x_0$, then

$$f_n(\tilde{x}_n, 0) = f(\tilde{x}_n) + |\tilde{x}_n - x_0|^4 \geq f(x_0) + |\tilde{x}_n - x_0|^4 > f(x_0) = 0.$$

On the other hand, we should have $f_n(\tilde{x}_n, 0) \leq f_n(\tilde{x}_n, 1) = 0 = f(x_0)$. This contradiction proves that $\tilde{x}_n = x_0$. In this case,

$$f_n(\tilde{x}_n, \tilde{\chi}_n) = \tilde{f}(x_0) = 0,$$

and, therefore, the minimum value to the $n-$problem would be equal to zero. Consequently, the point $(\hat{x}_n, \hat{\chi}_n) = (\tilde{x}_n, 1)$ also yields a solution to the $n-$problem because this point satisfies all the constraints of $n-$problem and $f_n(\tilde{x}_n, 1) = 0$. In addition, the last coordinate of this point is positive. Therefore, for any n, there exists a solution to the $n-$problem $(\hat{x}_n, \hat{\chi}_n)$ such that $\hat{\chi}_n > 0$. In what follows, only such solutions will be considered.

From (15), we conclude that the sequence $\{\hat{\chi}_n\}$ is bounded, which, together with $f_n(\hat{x}_n, \hat{\chi}_n) \leq 0$, implies $\tilde{f}(\hat{x}_n) \leq \text{const}|\tilde{f}(x_n)|$. Since $\tilde{f}(x_n) \to f(x_0) = 0$ as $n \to \infty$, we conclude that, for any limit point $\hat{x}$ of the sequence $\{\hat{x}_n\}$, we have $\tilde{f}(\hat{x}) \leq 0$, $\hat{x} \in C$, $F_1(\hat{x}) \leq 0$, $F_2(\hat{x}) = 0$. However, since $f(\hat{x}) \geq 0$, we have that $\hat{x} = x_0$ and, thus, that $\hat{x}_n \to x_0$ as $n \to \infty$.

Let us apply the necessary conditions obtained in the first stage to the solution of the $n-$problem. According to the observation made above, we can, for all sufficiently large n, choose a solution $(\hat{x}_n, \hat{\chi}_n)$ to the $n-$problem such that all inequalities (14) are strict. Therefore, since we have to deal with necessary conditions of local character, we shall ignore constraints (14). By taking this into account, we may write the Lagrangian for the $n-$problem as

$$\widetilde{\mathcal{L}}_n(x,\chi,\lambda) = \lambda_0 f_n(x,\chi) + \langle\lambda_1, F_1(x) - \chi F_1(x_n)\rangle + \langle\lambda_2, F_2(x) - \chi F_2(x_n)\rangle.$$

By virtue of the results obtained at stage I, there exists a λ_n and a linear subspace

$\Pi_n \subseteq \mathcal{I}$ such that

$$|\lambda_n| = 1, \quad \lambda_{0,n} \geq 0, \quad \lambda_{1,n} \geq 0, \tag{16}$$
$$\lambda_{0,n}\tilde{f}(x_n) + \langle \lambda_{1,n}, F_1(x_n)\rangle + \langle \lambda_{2,n}, F_2(x_n)\rangle = \lambda_{0,n}\gamma'(\hat{\chi}_n) \geq 0, \tag{17}$$
$$\frac{\partial \widetilde{\mathcal{L}}_n}{\partial x}(\hat{x}_n, \hat{\chi}_n, \lambda_n) \in -N(\hat{x}_n, C), \tag{18}$$
$$\Pi_n \subseteq \operatorname{Ker} \frac{\partial F}{\partial x}(\hat{x}_n), \tag{19}$$
$$\frac{\partial^2 \widetilde{\mathcal{L}}_n}{\partial x^2}(\hat{x}_n, \hat{\chi}_n, \lambda_n)[\xi, \xi] \geq 0 \quad \forall \xi \in \Pi_n, \quad \operatorname{codim}_{\mathcal{I}} \Pi_n \leq k. \tag{20}$$

Note that (17) is equivalent to the equality $\dfrac{\partial \widetilde{\mathcal{L}}_n}{\partial \chi} = 0$. In addition, we have omitted the second order variation with respect to χ in (20) because we do not need it.

By extracting a subsequence, we obtain some unit vector λ as the limit of the sequence λ_n. By taking the limit as $n \to \infty$ in (16) and (18), we conclude that $\lambda \in \Lambda$ (again, by using the upper semicontinuity of the set-valued mapping $N(\cdot, C)$).

Let us prove that $\lambda \in \Lambda_a$. By theorem , there exists a subspace $\Pi \subseteq \mathcal{I}$ such that $\operatorname{codim}_{\mathcal{I}} \Pi \leq k$, $\Pi \subseteq \operatorname{Ls}\{\Pi_n\}$. Take an arbitrary $b \in \Pi$. There exists a sequence $\{b_n\}$ such that, by extracting a subsequence, we obtain

$$b_n \in \Pi_n, \ \forall n, \text{ and } b_n \to b \text{ as } n \to \infty.$$

By virtue of (19), we have $\dfrac{\partial F}{\partial x}(\hat{x}_n)b_n = 0$, $\forall n$, and thus, $\dfrac{\partial F}{\partial x}(x_0)b = 0$. Therefore, $\Pi \subseteq \operatorname{Ker} \dfrac{\partial F}{\partial x}(x_0)$.

Similarly, from (20) we have that $\dfrac{\partial^2 \mathcal{L}}{\partial x^2}(x_0, \lambda)[b, b] \geq 0$, $\forall b \in \Pi$, and thus $\lambda \in \Lambda_a$.

Let us return to inequality (17). We have that $h \in \mathcal{K}$ implies that

$$\frac{\partial F_1}{\partial x}(x_0)h \leq 0, \quad \frac{\partial F_2}{\partial x}(x_0)h = 0, \quad \langle \frac{\partial f}{\partial x}(x_0), h\rangle \leq 0.$$

Hence, by expanding the mapping F_2 at the point x_0 up to the terms of the second-order, and by taking into account (13), we obtain

$$\begin{aligned} F_2(x_n) &= F_2(x_n) - F_2(x_0) \\ &= \frac{\partial F_2}{\partial x}(x_0)(x_n - x_0) + \frac{1}{2}\frac{\partial^2 F_2}{\partial x^2}(x_0)[(x_n - x_0)]^2 + o(|x_n - x_0|^2) \\ &= \varepsilon_n \frac{\partial F_2}{\partial x}(x_0)h + \frac{1}{2}\varepsilon_n^2 \frac{\partial F_2}{\partial x}(x_0)w + \frac{1}{2}\varepsilon_n^2 \frac{\partial^2 F_2}{\partial x^2}(x_0)[h, h] + o(\varepsilon_n^2). \end{aligned}$$

Therefore,

$$\langle \lambda_{2,n}, F_2(x_n)\rangle = \frac{1}{2}\varepsilon_n^2 \left(\langle \lambda_{2,n}, \frac{\partial^2 F_2}{\partial x^2}(x_0)[h, h]\rangle + \langle \lambda_{2,n}, \frac{\partial F_2}{\partial x}(x_0)w\rangle \right) + o(\varepsilon_n^2).$$

Analogously, for the mapping F_1, (by taking into account that $\langle\lambda_{1,n}, \frac{\partial F_1}{\partial x}(x_0)h\rangle \leq 0$) we obtain

$$\langle\lambda_{1,n}, F_1(x_n)\rangle \leq \frac{1}{2}\varepsilon_n^2\left(\langle\lambda_{1,n}, \frac{\partial^2 F_1}{\partial x^2}(x_0)[h,h]\rangle + \langle\lambda_{1,n}, \frac{\partial F_1}{\partial x}(x_0)w\rangle\right) + o(\varepsilon_n^2).$$

Similarly, for the function f (by using the fact that $\langle\frac{\partial f}{\partial x}(x_0), h\rangle \leq 0$), we obtain

$$\langle\lambda_{0,n}, f(x_n)\rangle \leq \frac{1}{2}\varepsilon_n^2\left(\langle\lambda_{0,n}, \frac{\partial^2 f}{\partial x^2}(x_0)[h,h]\rangle + \langle\lambda_{0,n}, \frac{\partial f}{\partial x}(x_0)w\rangle\right) + o(\varepsilon_n^2).$$

By substituting these three relations into inequality (17), by dividing by ε^2, and, then, by passing onto the limit as $n \to \infty$, we obtain (5). Thus, for each $w \in O_C^2(x_0, h)$, we proved (5).

STAGE III. We now prove (5) in its full generality (i.e., for $\dim X = \infty$). Fix an arbitrary vector $h \in \mathcal{K}(x_0)$ and take any $w \in O_C^2(x_0, h)$. We denote by $\widetilde{\mathcal{M}}$ the set of all finite-dimensional subspaces $M \in \mathcal{M}$ such that $h \in M$, $F'(x_0)(M) = \mathrm{Im}F'(x_0)$, and $w \in O^2_{C\cap M}(x_0, h)$. We take an arbitrary $M \in \widetilde{\mathcal{M}}$ and consider the problem obtained from the initial problem by replacing X by M. For this finite-dimensional problem, according to what was proved at stage II, there exist Lagrange multipliers $\lambda_M = (\lambda_M^0, \lambda_{1,M}, \lambda_{2,M})$ such that

$$\begin{aligned}
&\lambda_M^0 \geq 0, \quad \lambda_{1,M} \geq 0, \quad |\lambda_M| = 1,\\
&\frac{\partial \mathcal{L}}{\partial x}(x_0, \lambda_M) \in -N^M(x_0, C),\\
&\frac{\partial^2 \mathcal{L}}{\partial x^2}(x_0, \lambda_M)[h]^2 + \langle\frac{\partial \mathcal{L}}{\partial x}(x_0, \lambda), w\rangle \geq 0,
\end{aligned}$$

and the index of the quadratic form $\frac{\partial^2 \mathcal{L}}{\partial x^2}(x_0, \lambda_M)$ on the subspace $M \cap \mathrm{Ker}F'(x_0)$ is not greater than $\mathrm{codim}_{\mathcal{J}}(\mathrm{Im}F'(x_0))$.

The set of such vectors λ_M is denoted by $\Lambda_a(M, h, w)$. According to the discussion above this set is nonempty and it is easy to see that it is closed for any $M \in \widetilde{\mathcal{M}}$. Moreover, for arbitrary $M_1, \ldots, M_n \in \widetilde{\mathcal{M}}$, we obviously have

$$\bigcap_{i=1}^{n} \Lambda_a(M_i, h, w) \supseteq \Lambda_a(M_1 + \ldots + M_n, h, w) \neq \emptyset.$$

Consequently, the system of sets $\Lambda_a(M, h, w)$, $M \in \widetilde{\mathcal{M}}$, is centered. Therefore, by the compactness of the unit sphere in $\mathbb{R}^{k+1}$, the intersection $\bigcap_{M\in\widetilde{\mathcal{M}}} \Lambda_a(M, h, w)$ is not empty.

We take an arbitrary vector $\lambda \in \bigcap_{M\in\widetilde{\mathcal{M}}} \Lambda_a(M, h, w)$. Obviously, the relations

$$\lambda \in \Lambda \text{ and } \frac{\partial^2 \mathcal{L}}{\partial x^2}(x_0, \lambda)[h]^2 + \langle\frac{\partial \mathcal{L}}{\partial x}(x_0, \lambda), w\rangle \geq 0$$

hold for this vector. Now, to show that $\lambda \in \Lambda_a$, it suffices to prove the existence of a subspace Π satisfying (3).

Indeed, the maximum dimension of a subspaces in $\mathcal{I} \cap \operatorname{Ker} F'(x_0)$ on which the quadratic form $\frac{\partial^2 \mathcal{L}}{\partial x^2}(x_0,\lambda)$ is negative definite does not exceed $(k - \dim(\operatorname{Im} F'(x_0)))$. This assertion is implied by the fact that $\lambda \in \Lambda_a(M,h,w)$ for any finite-dimensional subspace $M \in \widetilde{\mathcal{M}}$. The existence of a subspace Π that satisfies (3) now follows from the following lemma (Lemma 3.1 from Ch.1 in [1]):

Let q be a finite quadratic form defined on a vector space X. The index of q is equal to the codimension of the subspace of minimum codimension on which this form is nonnegative definite.

Thus, (5) is proved.

STAGE IV. Let us prove (4). Fix an arbitrary vector $h \in \mathcal{K}(x_0)$, and take any convex set $\mathcal{T}(h) \subseteq O^2_C(x_0,h)$. Note that, if $\mathcal{T}(h) = \emptyset$, then $\sigma(\cdot,\mathcal{T}(h)) = -\infty$, and (4) holds trivially. Therefore, we assume that the set $\mathcal{T}(h)$ is nonempty. By virtue of the results obtained in Stage III, relation (5) holds for each $w \in \mathcal{T}(h)$, and, thus, we may write the following inequality

$$\inf_{w\in\mathcal{T}(h)} \max_{\lambda\in\Lambda_a} \left(\frac{\partial^2 \mathcal{L}}{\partial x^2}(x_0,\lambda)[h,h] + \langle \frac{\partial \mathcal{L}}{\partial x}(x_0,\lambda), w\rangle \right) \geq 0.$$

Since the set $\mathcal{T}(h) \subseteq X$ is convex, the set $\Lambda_a \subseteq \mathbb{R}^{k+1}$ is convex and compact, and the function $\frac{\partial^2 \mathcal{L}}{\partial x^2}(x_0,\lambda)[h,h] + \langle \frac{\partial \mathcal{L}}{\partial x}(x_0,\lambda), w\rangle$ is bilinear with respect to the variables $w \in X$ and $\lambda \in \mathbb{R}^{k+1}$, we may use proposition (see Appendix below) to conclude that the inequality above holds when the left hand side is replaced by

$$\max_{\lambda\in\Lambda_a} \inf_{w\in\mathcal{T}(h)} \left(\frac{\partial^2 \mathcal{L}}{\partial x^2}(x_0,\lambda)[h,h] + \langle \frac{\partial \mathcal{L}}{\partial x}(x_0,\lambda), w\rangle \right),$$

and, thus,

$$\max_{\lambda\in\Lambda_a} \left(\frac{\partial^2 \mathcal{L}}{\partial x^2}(x_0,\lambda)[h,h] + \inf_{w\in\mathcal{T}(h)} \left(\langle \frac{\partial \mathcal{L}}{\partial x}(x_0,\lambda), w\rangle \right) \right) \geq 0.$$

This and the arbitrariness of h imply (4). The theorem is proved. □

Appendix

Given a finite-dimensional Euclidean space Z (say $Z = \mathbb{R}^m$), a convex compact set $\tilde{Z} \subset Z$, a vector $a \in Z$, a convex subset $\mathcal{T} \subseteq X$, and a linear mapping $A : X \to Z$.

Let it be

$$\inf_{w\in\mathcal{T}} \max_{z\in\tilde{Z}} \langle Aw + a, z\rangle \geq 0. \tag{21}$$

Then

$$\max_{z\in\tilde{Z}} \inf_{w\in\mathcal{T}} \langle Aw + a, z\rangle \geq 0. \tag{22}$$

Proof. If $0 \in \tilde{Z}$, then (22) obviously holds. Therefore, suppose that $0 \notin \tilde{Z}$. Then, it is easy to see that cone$\tilde{Z}$ is closed. Here, we use the fact that the set $\tilde{Z}$ is compact. We consider a set $W = A\mathcal{T} + a$. Evidently, this set W is convex, and, from (21), we conclude that $W \bigcap int(\text{cone}\tilde{Z})^0 = \emptyset$. Hence, we can separate the sets W and $int(\text{cone}\tilde{Z})^0$. This means that there exists a vector $\bar{z} \in Z$ such that $\bar{z} \neq 0$, $\bar{z} \in (\text{cone}\tilde{Z})^{00}$, $\langle \bar{z}, z \rangle \geq 0$, $\forall z \in W$. But $(\text{cone}\tilde{Z})^{00} = \text{cone}\tilde{Z}$ because $(\text{cone}\tilde{Z})$ is closed. Hence, $\bar{z} \in \text{cone}\tilde{Z}$. Take $z_0 \in \tilde{Z}$, such that $z_0 = \alpha\bar{z}$, $\alpha > 0$. Then, $\langle Aw + a, z_0 \rangle \geq 0$, $\forall w \in \mathcal{T}$. The proposition is proved. □

References

[1] A. V. Arutyunov, *Optimality conditions: Abnormal and degenerate problems*, Kluwer Academic Publishers, Dordrecht, Boston, London, 2000.

[2] ———, *Necessary extremum conditions and a inverse function theorem without a priori normality assumptions*, Proceedings of the Steklov Institute of Mathematics **236** (2002), 25–36.

[3] A. V. Arutyunov, V. Jacimovic, and F. Pereira, *Second order necessary conditions for optimal impulsive control problems*, Journal of Dynamical and Control Systems **9** (2003), 131–153.

[4] J. F. Bonnans, R. Cominetti, and A. Shapiro, *Second order optimality conditions based on parabolic second order tangent sets*, SIAM J. Optimization **9** (1999), 466–492.

[5] R. Cominetti and J. Penot, *Tangent sets of order one and two to positive cones of some functional spaces*, Applied Mathematics and Optimization **36** (1997), 291–312.

[6] B. S. Mordukhovich, *Maximum principle in problems of time optimal control with nonsmooth constraints*, Appl. Math. Mech. **40** (1976), 960–969.

[7] ———, *Approximation methods in problems of optimatization and control*, Nauka, Moscow, 1988.

[8] ———, *Complete characterization of openness, metric regularity, and lipshitzian properties of multifunctions*, Transactions of the American Mathematical Society **340** (1993), 1–36.

[9] R. Rockafellar and R. Wets, *Variatitional analysis*, Springer, Berlin, 1998.

[10] A. Shapiro, *First- and second-order analysis of nonlinear semidefinite programs*, Mathematical Programming **77** (1997), 301–320.

[11] R. B. Vinter, *Optimal control*, Birkhauser, Boston, 2000.